

CLAIMS :

1. A method for transforming a heterogeneous compound document to a desired format based on a prescribed model comprising the steps of:
 - (i) specifying the components of the heterogeneous compound document as a hierarchical tree structure in multiple formats and specifying the input sources of information for each of the components as part of the model;
 - (ii) verifying and composing the heterogeneous compound document using a modeling language parser and a document composer by dynamically obtaining input information from the sources specified in the model and
 - (iii) converting the input compound document to the desired format using the document composer.
2. The method as claimed in claim 1, wherein the step of specifying components of heterogeneous compound document as a hierarchical tree structure is carried out by:

(i) identifying hierarchical structure of the compound document and determining the root nodes, internal nodes and leaf nodes;

(ii) creating a leaf element in the model description of the document for each leaf node and describing the leaf node in each of the required formats, with the proviso that for each value in the node, a stream start symbol, a stream descriptor, the value and stream end symbol are added to list of values corresponding to the given format in the leaf element, if input is to be obtained from a new source of data, otherwise the constant value specified by the leaf node is added to the list of values corresponding to the format in the leaf element;

(iii) creating an internal node element in the model of the document for each identified internal node and describing the internal node in each of the required formats for each child node of the node with the proviso that (a) if the input for the child node is to be obtained from a new source of data, a stream start symbol and a stream descriptor are added to the list of values corresponding to the given format in the internal node element; (b) if the child node is leaf node, the corresponding leaf element is added to the list of values corresponding to the given format and a stream end symbol is also added, if a stream start symbol was added in step (a); (c)

if the child node is an internal node, a reference to the corresponding internal node element is added to the list of values corresponding to the given format and a stream end symbol is also added, provided a stream start symbol was added in step (a);

(iv) creating a root element in the model of the document for each identified root node and describing a root node in each of the required formats for each child node of node with the proviso that (a) if the input too the child node is to be obtained from a new source of data, a stream start symbol and a stream descriptor are added to the list of values corresponding to the given format in the root element, b) if the child node is a leaf node, a reference to the corresponding leaf element is added to the list of values corresponding to the given format in the root element, and a stream end symbol is also added if a stream start symbol was added in step (a), (c) if the child node is an internal node, a reference to the corresponding internal node element of the model is added to the list of values corresponding to the given format of the root element, and a stream end symbol is also added if a stream start symbol was added in step (a).

3. The method as claimed in claim 1, wherein the step of verifying and composing of the heterogeneous compound document is carried out by:

- (i) providing a model of the compound document and the document format identifier as inputs;
- (ii) parsing the input model using a modeling language parser and updating the model database with parsed model information;
- (iii) obtaining a default input stream for the implementation of the method and pushing the same into a stream stack;
- (iv) considering each root element in the parsed model information and obtaining a value list corresponding to the input document format specified by the input document format identifier from the root element;
- (v) performing the following steps on the value list obtained in the previous step:
 - a) repeating step (v) with next value for a constant value which matches with the input values obtained from the stream on top of the stream stack;

- b) obtaining a value list corresponding to the internal node element or leaf element from model database and recursively repeating step (v) with a new list for a value which is a reference to an internal node element or a leaf element;
 - c) obtaining the next value which will be the stream descriptor for a value which is a stream start symbol and providing the stream descriptor to the stream module to obtain a new stream and pushing it into the stream stack;
 - d) popping the top most stream from the stream stack and using the current top most stream for further processing for a value which is a stream end symbol;
- (vi) popping all the streams from the stream stack except for the default input stream, resetting the default input stream and repeating step (iv) onwards with the next root element in the model, till all values in the value list match with the input and the root element contains the compound document information.

4. The method as claimed in claim 1, wherein the step of converting the heterogeneous compound document to the desired format specified in the prescribed model is carried out by:

- (i) providing a model of the compound document and the root element of the model matching with the input, and the output document format identifier as inputs;
- (ii) obtaining a value list corresponding to the output document format identifier from the root element;
- (iii) considering each value from the value list obtained and
 - (a) outputting the value if the value obtained is a constant;
 - (b) obtaining the value list corresponding to the output document format identifier from the internal element or leaf element, if the value is a reference to an internal node element or leaf element and recursively repeating step (iii) for the new value list.

5. A system for specifying, verifying, dynamically composing and transforming a heterogeneous compound document based on a

prescribed model, comprising a modeling language for depicting the structure of a compound document in one or more formats as a hierarchical tree structure and specifying the information source for each component of the document, providing a root element describing root nodes which are at the top of the structure, an internal node element representing the internal nodes of the compound document and a leaf element representing the leaf nodes of the bottom the structure; a modeling language parser for parsing the models created using the modeling language, analysing the given model, checking whether it confirms with the modeling language syntax and creating an internal representation of the model elements; a model database which stores information about each element defined in the model; a stream module for obtaining data dynamically from different sources specified in the stream specification in a model for composing compound document; and a document composer for composing the document in the format specified by the model using information from the model database and obtain input from multiple sources using the stream module and performing transformation on the composed document to convert it into any other format specified in the model.